



The Structure of Nucleic Acids

DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) are polymers of nucleotides linked in a chain through phosphodiester bonds. In biological systems, they serve as information-carrying molecules or, in the case of some RNA molecules, catalysts. This brief review will focus on aspects of structure of particular importance in manipulating DNA.

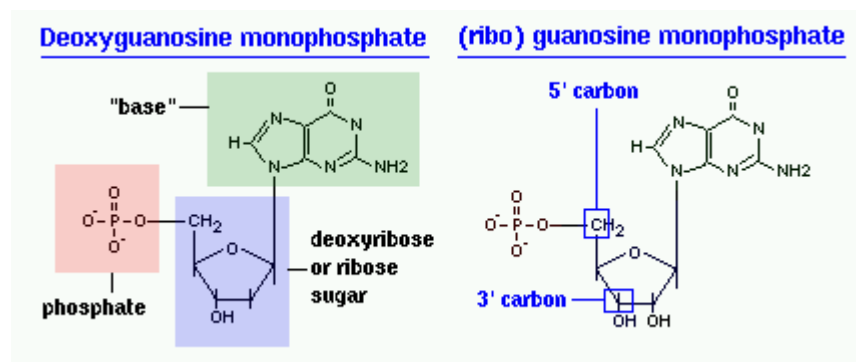
Bases, Nucleosides and Nucleotides

Nucleotides are the building blocks of all nucleic acids. Nucleotides have a distinctive structure composed of three components covalently bound together:

- **a nitrogen-containing "base"** - either a pyrimidine (one ring) or purine (two rings)
- **a 5-carbon sugar** - ribose or deoxyribose
- **a phosphate group**

The combination of a base and sugar is called a *nucleoside*. Nucleotides also exist in activated forms containing two or three phosphates, called nucleotide diphosphates or triphosphates. If the sugar in a nucleotide is deoxyribose, the nucleotide is called a deoxynucleotide; if the sugar is ribose, the term ribonucleotide is used.

The structure of a nucleotide is depicted below. The structure on the left - deoxyguanosine - depicts the base, sugar and phosphate moieties. In comparison, the structure on the right has an extra hydroxyl group on the 2' carbon of ribose, making it a ribonucleotide - riboguanosine or just guanosine.



In the right-hand figure, note also the 5' and 3' carbons on ribose (or deoxyribose) - understanding this concept and nomenclature is critical to understanding polarity of nucleic acids, as discussed below. The 5' carbon has an attached phosphate group, while the 3' carbon has a hydroxyl group.

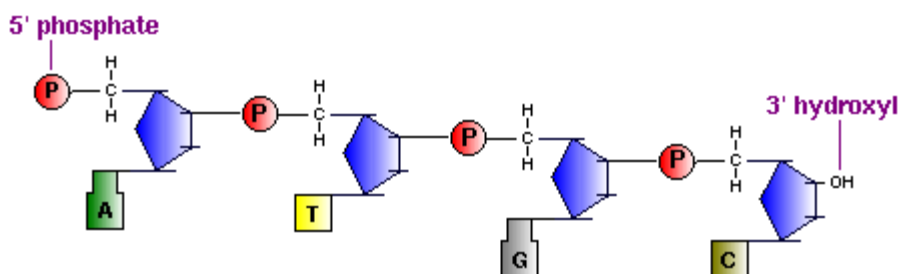
There are five common bases, and four are generally represented in either DNA or RNA. Those bases and their corresponding nucleosides are described in the following table:

Abbr.	Base	Nucleoside	Nucleic Acid
A	Adenine	deoxyadenosine	DNA
		adenosine	RNA
G	Guanine	deoxyguanosine	DNA
		guanosine	RNA
C	Cytosine	deoxycytidine	DNA
		cytidine	RNA
T	Thymine	deoxythymidine (thymidine)	DNA
U	Uracil	uridine	RNA

Another useful way to categorize nucleotide bases is as *purines* (A and G) versus *pyrimidines* (C, T and U). Although [committing this to memory](#) is often difficult, the importance is that in double-stranded nucleic acids, base pairs are always formed between a purine and a pyrimidine.

Nucleic Acids

DNA and RNA are synthesized in cells by DNA polymerases and RNA polymerases. Short fragments of nucleic acids also are commonly produced without enzymes by oligonucleotide synthesizers. In all cases, the process involves forming phosphodiester bonds between the 3' carbon of one nucleotide and the 5' carbon of another nucleotide. This leads to formation of the so-called "sugar-phosphate backbone", from which the bases project.

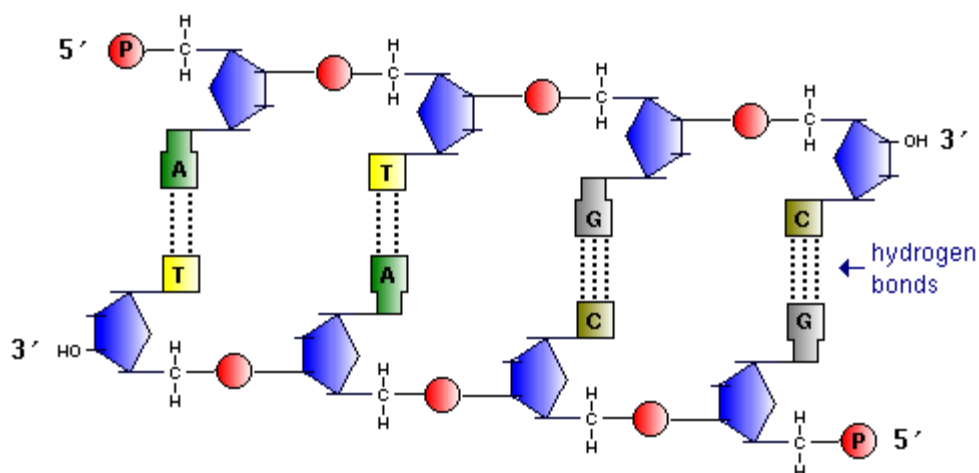


A key feature of all nucleic acids is that they have two distinctive ends: the 5' (5-prime) and 3' (3-prime) ends. This terminology refers to the 5' and 3' carbons on the sugar. For both DNA (shown above) and RNA, the 5' end bears a phosphate, and the 3' end a hydroxyl group.

Another important concept in nucleic acid structure is that DNA and RNA polymerases add nucleotides to the 3' end of the previously incorporated base. Another way to put this is that nucleic acids are synthesized in a 5' to 3' direction.

Base Pairing and Double Stranded Nucleic Acids

Most DNA exists in the famous form of a double helix, in which two linear strands of DNA are wound around one another. The major force promoting formation of this helix is **complementary base pairing**: A's form hydrogen bonds with T's (or U's in RNA), and G's form hydrogen bonds with C's. If we mix two ATGC's together, the following duplex will form:

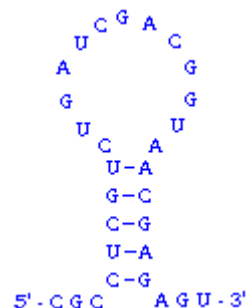


Examine the figure above and note two very important features:

- **The two strands of DNA are arranged antiparallel to one another:** viewed from left to right the "top" strand is aligned 5' to 3', while the "bottom" strand is aligned 3' to 5'. *This is always the case for duplex nucleic acids.*
- **G-C base pairs have 3 hydrogen bonds, whereas A-T base pairs have 2 hydrogen bonds:** one consequence of this disparity is that it takes more energy (e.g. a higher temperature) to disrupt GC-rich DNA than AT-rich DNA.

The figures above fail to impart any appreciation of the three-dimensional structure of DNA. This deficiency can be rectified to some extent by viewing and manipulating a [3-D model of duplex DNA](#).

What about double stranded RNA? RNAs are usually single stranded, but **many RNA molecules have secondary structure in which intramolecular loops are formed by complementary base pairing**. A simple example of this is shown in the figure to the right, and much more extensive and complex examples are known. Base pairing in RNA follows exactly the same principles



as with DNA: the two regions involved in duplex formation are antiparallel to one another, and the base pairs that form are A-U and G-C.